

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2003, 17:23:34 ; Search time 2961 Seconds
(without alignments)
4835.653 Million cell updates/sec

Title: US-09-745-506-37
Perfect score: 1799
Sequence: 1 MDKALSLSLNDFASLSFAE.....LENKINIILSETDRDPLQVY 350

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODER-frame+ p2n.model -DEV=tlh
-Q/cgn2.1/USPTO_spool/US09745506/runat_22082003_104401_7037/app-query.fasta.1.519
-DB-genemb1 -QFMT-fastap -SUFFIX-p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09745506.ecgn.1.1.2377 -runat_22082003_104401_7037 -NCPU=6 -ICPU=3
-NO MAP -LARGEDUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1799	100.0	1053	6	AX119075	AX119075 Sequence
2	1799	100.0	1387	9	AF060513	AF060513 Homo sapi
3	1799	100.0	1425	9	AF283538	AF283538 Homo sapi
4	1799	100.0	1440	9	BC007654	BC007654 Homo sapi
5	1799	100.0	1574	6	BD158389	BD158389 Homo sapi
6	1799	100.0	1574	9	AK023378	AK023378 Homo sapi
7	1799	100.0	1606	9	AB038949	AB038949 Homo sapi
8	1763	98.0	1579	9	AF182416	AF182416 Homo sapi
9	1715	95.3	1353	9	HSMB05522	HSMB05522 Homo sapi
10	1575	87.5	1836	10	AF284439	AF284439 Mus muscu
11	976.5	54.3	55374	9	AL645474	AL645474 Human DNA
12	826.5	45.9	796	6	BD149184	BD149184 primer fo
13	810	45.0	1328	3	AK114307	AK114307 Clona int
14	737	41.0	170586	2	AC037455	AC037455 Homo sapi
15	737	41.0	190508	9	AC005037	AC005037 Homo sapi
16	737	41.0	198250	2	AC093681	AC093681 Homo sapi
17	601	33.4	231600	2	AC130779	AC130779 Rattus no
18	585	32.5	149819	2	AC121091	AC121091 Mus muscu
19	585	32.5	234976	2	AC118698	AC118698 Mus muscu
20	578	32.1	1036	3	AY069556	AY069556 Drosophi1
21	578	32.1	70667	2	AC017306	AC017306 Drosophi1
22	578	32.1	165839	3	AC092398	AC092398 Drosophi1
23	578	32.1	273414	3	AE003650	AE003650 Drosophi1
24	578	32.1	299935	3	DROSADH09	AE003415 Drosophi1
25	398	22.1	155127	2	AL929495	AL929495 Dario rer
26	386	21.5	249	6	BD049005	BD049005 Sequence
27	384.5	21.4	1187	8	SCVGL221C	272743 S. cerevisia
28	358.5	19.9	12320	1	AE013129	AE013129 Thermomana
29	353	19.6	901	8	SCU52042	US2042 Saccharomyc
30	351	19.5	28324	8	SPPC126	AL034490 S. pombe c
31	317.5	17.6	300050	1	AP004599	AP004599 Oceanobac
32	312	17.3	300029	1	AE016748	AE016748 Staphyloc
33	305	17.0	3041	1	AF270328	AF270328 Staphyloc
34	305	17.0	3041	6	AX145646	AX145646 Sequence
35	303	16.8	302173	1	AE016951	AE016951 Enterococ
36	296	16.5	301550	1	AP003192	AP003192 Clostridi1
37	288.5	16.0	281017	2	AX143427	AX143427 Sequence
38	288.5	16.0	281017	6	AC123462	AC123462 Rattus no
39	288.5	16.0	290029	1	AE017038	AE017038 Bacillus
40	288	16.0	3441	1	AF269767	AF269767 Staphyloc
41	288	16.0	3441	6	AX145085	AX145085 Sequence
42	285.5	15.9	218470	1	BSUB0013	299116 Bacillus su
43	284.5	15.8	302422	1	AE017011	AE017011 Bacillus
44	281.5	15.6	301439	1	AE015943	AE015943 Clostridi1
45	279	15.5	333750	1	AP004827	AP004827 Staphyloc

RESULT 1

ALIGNMENTS

AX119075
LOCUS AX119075 1053 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 239 from Patent WO0129221.
ACCESSION AX119075
VERSION AX119075.1 GI:14036029
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 Conklin, D.C. and Yee, D.P.
Proteins and polynucleotides encoding them
Patent: WO 0129221-A 239 26-APR-2001;
ZymoGenetics, Inc. (US)
FEATURES
source
1. 1053
location/Qualifiers
CDS
1. 1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC38611.1"
/db_xref="GI:14036030"
/translation="MDKLKLSLNDPFLASISFAESWDNVGLVPEPSPPHYVNTLEFLN
DLREVEVEYLQKRAKDLISYHPDIFPMKRTWNTWKERLVIALENRGVYSPTA
YDAPOGVNMMIAKAGACTSRPIHSPKAPYPTGKNRHFVNNYDODLVKSAVK
GIDGVSATSFASRTGNEQRTINLNCQKALMOYVDVFSRKOLYOKTEILSKPL
LHTGMGRCLTDESIVSLATMTDRIKRHLKXSHIRLALGVSTLSQKVVAVLGGSGS
SYLOGVERADLYLTREMSHHDLDLAASGIVNIIICEHSTNGEFLSDLRMDLSLENK
INITLSETRDPLQV"
BASE COUNT 279 a 253 c 247 g 274 t
ORIGIN
Alignment Scores:
Pred. No.: 4,43e-146 Length: 1053
Score: 1799.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-745-506-37 (1-350) x AX119075 (1-1053)
QY 1 MetaspLeuYsAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB 1 ATGATTTTGAAGGGCTCTCTTCTTCTTGAATGACCTTGATCCCTCGTTGCTGAG 60
QY 21 SerTTPAspAsnValGlyLeuLeuValGluProSerProProHisThValAsnThre 40
DB 61 AGTGGGACAAATGTTGATCTGTGTGACCAAGCCACACACATACATGTAATCACACTC 120
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLysLysAlaSp 60
DB 121 TTCTGACCAATGACCTGACTGAGCACTGATGGAGGAGGCTCTCAAAAGAGCAAGC 180
QY 61 LeuLeuSerTyrHisProProIlePheArgProMetLysArgIleThrTTPAsnThr 80
DB 181 CTCATCTCTCTCCATCCGCGCTATCTTCGAGACCACGAGGCGATTAACCTGGAACACA 240
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
DB 241 TGGAGAGGAGCGCTGGTATCCGGCTCTGGAGAACAGATCGGTATCTACTCTCTCAT 300
QY 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTTPLeuAlaLysGlyLeuGlyAla 120
DB 301 ACAAGCTATGATGCTGCGCCCGACGAGGCTCAACAACATGTTGGCTTAAGAGGCTTGAGCT 360
QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGluLysHis 140
DB 361 TGTACCTCCAGGCCCATGATCTCTCCAAAGCTCCCAACTACCTTACAGAGGAGAACAC 420

QY 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
DB 421 CGAGTAGAATTCACAGTTAACTACACCCAAACCTGGACAAAGTATGCTGCGATGAAA 480
QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluThr 180
DB 481 GGAATGACGCTGTTCTGTCACCTCTTTTCTGCTGAGACTGGTATAGAGAACAAACA 540
QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGluValValAspPheLeuSerArg 200
DB 541 CGGATTAATCTGAAATGACTACAGAGAGCTTGTATGACGAGTGTAGATTTCTTCCCGG 600
QY 201 AsnLysGlnLeuTyrGlnLysThrGlnLysLeuSerLeuGluLysProLeuLeuHis 220
DB 601 AACAAACAACTTTATCAGAACAGCGAAATTCGTACAGTGGAGAACCTTTGCTTCAT 660
QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
DB 661 ACTGGAATGGACCGTTATGACACACTGATGAATCTGCTCTCCGACACCATGATTCAT 720
QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValArgThr 260
DB 721 CGAATTAATAAGACACCTAAACCTAATCTCATATTCGCTTAGCCCTGGGGTGGAGAAC 780
QY 261 LeuGluSerGlnValLysValValAlaLeuGluAlaGlySerGlySerSerValLeuGln 280
DB 781 TTAAGAGTCTCAAGCAAGAGTGTGGCCCTGTGCTGGTTGTGGAGAGCGTCTCGAG 840
QY 281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisAspThrLeuAspAla 300
DB 841 GGTGTGAGGCTGACCTTACTCTACACAGCTAGATGCCATCATGATCTTGGATGCT 900
QY 301 AlaSerGlnGlyIleAsnValIleLeuGluSerGluHisSerAsnThrGluArgGlyPheLeu 320
DB 901 GCTTCCCAAGGAATAAATGTCATCTCTGTGAACACACACACTGAACGAGGCTTCTT 960
QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer 340
DB 961 TCTGACCTTCAGATATCTGATTTCTACCTTGGAGAAATTAATTAATTAATCTTATCA 1020
QY 341 GluThrAspArgAspProLeuGlnValVal 350
DB 1021 GAGACTGACAGGAGACCTTCCAGGTGCTA 1050
RESULT 2
AF060513 1387 bp mRNA linear PRI 02-JAN-2001
LOCUS AF060513
DEFINITION Homo sapiens clone 016d06 My018 protein mRNA, complete cds.
ACCESSION AF060513
VERSION AF060513.1 GI:12001975
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 1387)
AUTHORS Mao, Y.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
source
1. 1387
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="016d06"
/tissue_type="brain"
/dev_stage="fetus"
/dev_stage="fetus"
/dev_stage="fetus"
/codon_start=1
/product="My018 protein"
/protein_id="AAG43131.1"

/db_xref="GI:12001976"
 /translation="MDIKALLSLNDFASLSPAESMDNGLVPEPSPPTVNTLFLTN
 DLTEFMEVLEOKKADILISYRPIERPKRITRNRKRLVIFALENVGIXSPETA
 YPAAGVYRNMLAKGLGACTSRPIHPSKAPNPTBGNHVEVNTTODLDKMAVK
 GIDGVSVTSFSAKNEQTRINLNTOTALMOVDLFSRNOLVOKTILSELKLL
 LATGVRICLTDESVSLAMIDIRIKRLHSLRLAGVRILKLEISQVKVILACAGSGS
 SVLQGVADLYLTGEMSHHDTLDAASOGINVLICEHSNTERGFI.SDLRMLDSHLENK
 INILSETRDLPLVY"
 BASE COUNT 378 a 323 c 309 g 377 t
 ORIGIN

Alignment Scores:

Pred. No.: 6,2e-146 Length: 1387
 Score: 1799.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x AF060513 (1-1387)

OY 1 MetaSPLeuLYaAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 82 ATGGATTGTAAGGCTCTCTTCTTCTGGAATGACTTTCATCCCTCTCTTGGCTGAG 141
 OY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
 DB 142 AGTTGGGACAAATGTTGGATTACTGTTGGAGCAAGCCCAACATCTGTAATAATACACTC 201
 OY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluInLysAlaAsp 60
 DB 202 TTCTGACCAATGACTGACTGAGGAAGATGATGGAGAGTGTGCAAAAAGAGCAGAC 261
 OY 61 LeuIleLeuSerTyrHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80
 DB 262 CTCATTCTCTCTACCATCGGCTATCTCCGACCATGAGGCGCATTAACCTGGACACACA 321
 OY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
 DB 322 TGGMAGAGCGCTGATCCGGGCTCTGGAGAACAGAGTCGTAATCTCTCTCAT 381
 OY 101 ThrAlaTyrAspAlaIleProGluGluValAsnAsnTrpLeuAlaLysGlyLeuGluAla 120
 DB 382 ACAGCCTATGATGCTCGCCGAGGCGTCACAACTGTTGGCTAAAGGCTTGGAGCT 441
 OY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGluAsnHis 140
 DB 442 TGTACTCTCAGGCCCTATCATCTCTCAAGCTCCCACTACCTCAGAGGGAAACACAC 501
 OY 141 ArgValGluPheAsnValAsnTyrThrGluAspLeuAspLysValMetSerAlaValLys 160
 DB 502 CGAGTAGAATTCACACTTAACACCAAGACCTGGACAAAGCTCATGCTTCGAGTGA 561
 OY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluInThr 180
 DB 562 GGAATTTGAGGGTCTCTCTCTCTCTTCTGCTAGACACTGGTAATAGGAACAACA 621
 OY 181 ArgIleAsnLeuAsnCysThrGluLysAlaLeuMetGluValValAspPheLeuSerArg 200
 DB 622 CGGATTAATCTGATGTACTCAGAAAGGCTTGATGAGTGTAGATTCTTCTCCCG 681
 OY 201 AsnLysGluLeuTyrGluInLysThrGluIleLeuSerLeuGluLysProLeuLeuLeuHis 220
 DB 682 AACAAACACTTATTCAGAAAGGCAATTTCTGTCACTGAGAGAGCTTTCCTTCTACAT 741
 OY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
 DB 742 ACTGGATGGAAGCGGTTATGACACACTGGATGATCTGTCTCCCTGCAACATGATGAT 801
 OY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGluValGlyArgThr 260
 DB 802 CGAATTAAGAACACCTAAAACTATCTCATTTCTGCTTGGCCCTTGGGGTGGAGAACCC 861

OY 261 LeuGluSerGluValLysValAlaLeuCysAlaGlySerGlySerValLeuGluIn 280
 DB 862 TTGAGATCTCAAGTCAAAGTCGGGCCCTGTGCTGCTTCTGGAGCAGCGCTTCGACAG 921
 OY 281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300
 DB 922 GGTGTGAGGCTGACCTTAACTCACTCAGAGGATGATGCCATCATGATATTGATGCT 981
 OY 301 AlaSerGluGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGlyPheLeu 320
 DB 982 GCTTCCCAAGGATTAATGATCATCTCTGGAACACAGCAACCTGACAGCGCTTCTT 1041
 OY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer 340
 DB 1042 TCTGACCTTCGACATATGCTGATCTCTCACTTGGAAATAGATTAATATTATCTATCA 1101
 OY 341 GluThrAspArgAspProLeuGluValVal 350
 DB 1102 GAGACTGACAGGAGCCCTCTTCAGTGTGTA 1131

RESULT 3

AF283538 1425 bp mRNA linear PRI 16-JAN-2001
 LOCUS Homo sapiens NIF3L1 protein mRNA, complete cds.
 DEFINITION
 ACCESSION AF283538
 VERSION AF283538.1 GI:12006402
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1425)
 Tascou,S., Uedelhoven,J., Dikens,C., Mayernia,K., Engel,W. and
 Burfelind,P.
 Isolation and characterization of a novel human gene, NIF3L1, and
 its mouse ortholog, Nif3l1, highly conserved from bacteria to
 mammals

AUTHORS

Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany

TITLE

cytochrome. Cell Genet. 90 (3-4), 330-336 (2000)

JOURNAL

Medline 20573864
 PUBMED 11124544

REFERENCES

2 (bases 1 to 1425)
 Tascou,S., Burfelind,P. and Engel,W.
 Direct Submission
 Submitted (29-JUN-2000) Institute for Human Genetics, University of
 Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany

FEATURES

source
 1..1425
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 47..1180
 /codon_start=1
 /product="NIF3L1 protein"
 /protein_id="AA044846.1"
 /db_xref="GI:12006403"
 /translation="MSSGVPRPVPTVRPVDLSICNSSRSPMDLALLSLNDFASL
 PASHMDNGLVPEPSPPTVNTLFLNDLTERVMEVLEOKKADILISYRPIERPKR
 ITWTKERLVIRALENRVITSPHATDGVNMLAKGLGACTSRPIHPSKAPN
 MYPLEGHVRHFNVTNYTODLDKWSAVKSIDGVSVTSFSAKNEQTRINLNTOTAL
 MOVDLFSRNOLVOKTILSELKPLLLATGVRILKLEISQVKVILACAGSGS
 HIRLAAGVRLTESQYKVAALCAGSGSSVLAGVADLYLTGEMSHHDTLDAASOGIN
 ILCEHSNTERGFI.SDLRMLDSHLENKINILSETRDLPLVY"

CDS

BASE COUNT 385 a 331 c 314 g 395 t
 ORIGIN

Alignment Scores:

Pred. No.: 6,41e-146 Length: 1425
 Score: 1799.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x AF283538 (1-1425)

Oy 1 MetasplesinAlaLeuSerSerLeuAsnppheAlaSerLeuSerPheAlaGlu 20
 DB 128 ATGATTTGAAAGGCTCTCTCTCTCTCTGATGATCTTGCATCTCTCTCTCTGAG 187
 Oy 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProPonHisThrValAsnThrLeu 40
 DB 188 AGTTGGACAAATGTTGATTAAGTGTGGAACCAACCAACCAACCAATGTAATAATACACTC 247
 Oy 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnIlyAsp 60
 DB 248 TTCTGTACCAATGACCTGACTGAGAGAGTGTGAGAGAGTGTGCAAAAAGAGGAGAC 307
 Oy 61 LeuLeuSerSerTrpHisProPheArpPheArpProMetIlyAspGileThrTrpAsnThr 80
 DB 308 CTCATTTCTCTCTACCATCCGCTATCTTCGACACCATGAAGCCATTAACCTGGAACACA 367
 Oy 81 TrpIysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTrpSerProHis 100
 DB 368 TGGAAAGAGCGCTGGTGTATCCGGGCTGTGGAAGAGAGTGTATCTCTCTCAT 427
 Oy 101 ThrIleArgAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaIlyGlyLeuGlyAla 120
 DB 428 ACACCTTATGATGCTGGGCCCCAGGGGCTCAACAACTGCTGCTAAAGGCTTGAGCT 487
 Oy 121 CysThrSerArgProIleHisProSerIlyAspProAsnTrpProThrGluGlyAsnHis 140
 DB 488 TGTACTGTACAGCCCATATACCTCTTCCAAAGCTCCCAACTACCTACAGAGGAGAAACAC 547
 Oy 141 ArgValGluPheAsnValAsnTrpThrGlnAspLeuAspIlyValMetSerAlaValIly 160
 DB 548 CGAGTAGAATTAACAGTTAACTACACCAACCACTGTGCAAAAGTATCTGCACTGAGTAAA 607
 Oy 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnIleThr 180
 DB 608 GGAATTGACGGTGTCTCTGCTACTTCTTTCTGCTAGACTGTGTAAGAGAGAACAAACA 667
 Oy 181 ArgIleAsnLeuAsnCysThrGlnIlyAlaLeuMetGlnValValAspPheLeuSerArg 200
 DB 668 CGGATTAATCTGAATTTGACTACAGAAAGCTTGAATGACAGTGTGATTTCTTCCCGG 727
 Oy 201 AsnIysGlnLeuTrpGlnIlyThrGluLeuSerLeuGluIlyProLeuLeuHis 220
 DB 728 AACAAACAACTTTATCAAGAACAGGAATTTCTGTCTACAGGAACCTTTGCTCTCAT 787
 Oy 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
 DB 788 ACTGGAATGGAGCGTTATGCACACTGATGATCTCTCCCTGGCAACATGATTTGAT 847
 Oy 241 ArgIleIysArgHisLeuIlyLeuSerHisIleArgLeuAlaLeuGlyValIlyArgThr 260
 DB 848 CGAATTAAGAACACCTAAACTATCTATATTCCTTAAGCCCTTGGGGTGGAGAAC 907
 Oy 261 LeuGluSerGlnValIlyValAlaLeuCysAlaGlySerGlySerSerValLeuGln 280
 DB 908 TTAGAGTGTCAAGTAAAGCTGTGGCTGTGTCTGTCTGTGGAGAGCGCTTGTGAG 967
 Oy 281 GlyValGluAlaAspLeuTrpLeuThrGlyGluMetSerHisAspThrLeuAspAla 300
 DB 968 GGTGTGGAGGCTGACCTTACTCTCACAGTGAAGATGCTCCATCATGATCTTGGATGCT 1027
 Oy 301 AlaSerGlnGlyIlyLeuAsnValIleLeuCysGluHisSerAsnThrGluArgIlyPheLeu 320
 DB 1028 GCTTCCCAAGAAATAATGTATCTCTGTGTAACACAGAACACAGAACGAGCTTCTCT 1087
 Oy 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnIlyLeuGlnIleLeuSer 340
 DB 1088 TCTGACCTTCCGAGATATGCTGATTTCTCACTTGGAATTAAGATTAATATTTATCTCATCA 1147
 Oy 341 GluThrAspArgAspProLeuGlnValVal 350
 |||||||

DB 1148 GAGACTGACAGGAGCACTTTCAGGTGCTA 1177
 RESULT 4
 BC007654 1440 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, similar to NIF3 (Nrg1 interacting factor 3, S.pombe homolog)-like 1, clone MGC:2595 IMAGE:3346099, mRNA, complete cds.
 ACCESSION BC007654
 VERSION BC007654.1 GI:14043316
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1440)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/DMF
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland.
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nrl.nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stentrup, S., Thomas, P.J., Tlonsong, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IMAGE:3346099, 1 Column: 12
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10197631.
 FEATURES
 source
 1..1440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:2595 IMAGE:3346099"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NHL_MGC_20"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 37..1170
 /codon_start=1
 /product="Similar to NIF3 (Nrg1 interacting factor 3, S.pombe homolog)-like 1"
 /protein_id="AAH07654.1"
 /db_xref="GI:14043317"
 /translation="MTSCVAPVPTVTFVDSLONSSRSPMDIKALLSNDFAVSLFAESNDVGLVSPPTVTNTLPLTNDLTFEWEVYLQKADLISYHPPIPRMKRITWTWKRLVIRALENVGILSPHTAIDAPDQNNMLAGLACTSRPLHPKAPNYPTEGNHREVRNVNTDLDVYMSAVKIDSVYSFASARGNEQRIINLNSQKALMOVVDFLSRNKLOVKTETILSLERKLILHTGMRCLTDESVLSATMTDRKRLKLSHRLALGVGRILSOVKYVALCAGSGSSVLGVEDLVLTGESHHDPLDAASGGINVILCESNTERGFISDRMDSHLENKNIILSERDRPLQV"
 BASE COUNT 407 a 329 c 313 g 391 t
 ORIGIN
 Alignment Scores: 6.49e-146 length: 1440
 Pred. No.:

Score: 1799.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x BC007654 (1-1440)

QY 1 MetaspLeuAlaLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 118 ATGGATTGAAAGGCTCTCTTCTCTTGGAAAGATTCCTTCCCTCTCTTGGTGGAG 177
 QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
 DB 178 AGTTGGACAAATGTTGATTACTGCTGGAAACCAAGCCCAACATCTGTAATATACACTC 237
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLeuGluValAsp 60
 DB 238 TTCCTGACCACATGACCTGACAGCAAGATGAGAGAGTGTCTGCAAAAGAGCAGAC 297
 QY 61 LeuIleLeuSerThrHisProProIlePheArgPrometLeuArgIleThrTPAsnThr 80
 DB 298 CTCATTCTCTCTACCATCGGCTATCTTCCGACCATGAAAGCGCTTAACCTGGAAACACA 357
 QY 81 TPlySGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleThrSerProHis 100
 DB 358 TGGAAAGAGCGCTGGTGTATCCGGGCTCTGGAAACAGAGTGGTATCTACTCTCTCAT 417
 QY 101 ThrAlaTYRAspAlaAlaProGluGluValAsnAsnTPRLeuAlaGlyLeuGluAla 120
 DB 418 ACAGCTATGATGCTGCGCCCGCCAGGCGCTCAACACTGGTGGCTAAAGGCTTGGAGCT 477
 QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTPRProhrgIuGluAsnHis 140
 DB 478 TGTACTCTCCAGGCCCTACATCTCTTCCAAAGCTCCCACTACCTTCAAGAGGAAACAC 537
 QY 141 ArgValGluPheAsnValAsnTYRThrGlnAspLeuAspLysValMetSerAlaValLys 160
 DB 538 CGAGTGAATTCACAGCTTAACCTAACCAAGACCTGGACAAAGTCAATGCTGAGAGGAA 597
 QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyGluGluGluGlnThr 180
 DB 598 GGAATGACGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
 QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGluValValAspPheLeuSerArg 200
 DB 658 CGGATTAATCTGATCTGACTCAGAAAGCTTTCATGACGTGTATCTCTCTCTCTCT 717
 QY 201 AsnLysGluLeuTYRGLNLYSThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
 DB 718 AACCAACACTTATCAGAAAGAGGAAATTCGTCTGCTGAGAGAGGCTTGGCTTCTACAT 777
 QY 221 ThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetLeuAsp 240
 DB 778 ACTGGAATGGAGCGGTATGACACACTGATGATGCTGCTCTCTCTCTCTCTCTCTCT 837
 QY 241 ArgIleLeuArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGluValGlyArgThr 260
 DB 838 CCAATTAAGACACTTAATACTATCTCATATTCCTTACGCTTGGGGTGGGGAAGACC 897
 QY 261 LeuGluSerGluValLysValAlaLeuCysAlaGlySerGlySerSerValLeuGlu 280
 DB 898 TTAGAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 957
 QY 281 GlyValGluAlaAspLeuTYRLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300
 DB 958 GGTGTGGAGGCTGACCTTACTCAGAGGTGAGATGCCATCATATCTTGGTGGTGTCT 1017
 QY 301 AlaSerGluGlyIleAsnValIleLeuGlyGluHisSerAsnThrGluArgIlePheLeu 320
 DB 1018 GCTTCCCAAGATTAATATCTACTCTCTGAAACAGCAACAGCAACAGCAAGGCTTCTCT 1077
 QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleLeuSer 340

DB 1078 TCGACCTCGAGATATGCTGATTTCTCACTTGCAGAAATATATATATATATATATCA 1137
 QY 341 GluThrAspArgAspProLeuGluValVal 350
 DB 1138 GAGACTGACAGGAGGACCTCTTCAGTGTGTA 1167

RESULT 5

BD158389 1574 bp DNA linear PAT 17-JAN-2003

LOCUS Primer for synthesizing full-length cDNA and use thereof.

DEFINITION BD158389.1 GI:27864147

ACCESSION BD158389.1 GI:27864147

VERSION JP 2002191363-A/13232.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

JOURNAL Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

COMMENT Primer for synthesizing full-length cDNA and use thereof

OS Homo sapiens (human)

PN JP 2002191363-A/13232

PD 09-JUL-2002 JP 200280990

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUKI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH key

FEATURES Location/Qualifiers

source FT CDS (271)..(1320).

BASE COUNT 420 a 361 c 372 g 421 t

ORIGIN

Alignment Scores:

Pred. No.: 7.24e-146 Length: 1574

Score: 1799.00 Matches: 350

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-745-506-37 (1-350) x BD158389 (1-1574)

QY 1 MetaspLeuAlaLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 271 ATGGATTGAAAGGCTCTCTTCTCTTGGAAAGATTCCTTCCCTCTCTTGGTGGAG 330
 QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
 DB 331 AGTTGGACAAATGTTGATTACTGCTGGAAACCAAGCCCAACATCTGTAATATACACTC 390
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLeuGluValAsp 60
 DB 391 TTCCTGACCACATGACCTGACAGCAAGATGAGAGAGTGTCTGCAAAAGAGCAGAC 450
 QY 61 LeuIleLeuSerThrHisProProIlePheArgPrometLeuArgIleThrTPAsnThr 80
 DB 451 CTCATTCTCTCTACCATCGGCTATCTTCCGACCATGAAAGCGCTAACTGGAAACACA 510

QY 81 Trrpysgluargleuvalilleargalaleuclunsnargvalglylletyrserrprohis 100
 DB 511 TGGAGGAGCGCCGCGTGTATCCGGGCTTCGAGAACAGAGCGGTATCTTACTCTCCAT 570
 QY 101 ThrilatrraspaalaleaProglnglyValasnsntrpleaualalyglyleuclvala 120
 DB 571 ACACCCATGATGCTGCGCCGCCAGGGCTCAACAACCTGCTTGCTTAAGGCTTGAGACT 630
 QY 121 CysTrserArpProileHisProserLysAlaProasnTyrrProthrgluclLysnHis 140
 DB 631 TGTACCTCCAGGCCCATCATCTCTCCAAAGCTCCCAACTACCTACAGAGGAAACAC 690
 QY 141 ArgvalGluPheasnaValasnTyrrThrcLinspleuaplyValmeSerAlaVallys 160
 DB 691 CGAGTGAATTCACAGCTTAACTACACCCAAAGCTGGCAAGATCATCTGCAATGAAA 750
 QY 161 GlylleaspglyValserValThserPheSerAlaArghThrglyAsnGluGlnthr 180
 DB 751 GGAATGCGGTGTTCTGTCACCTCTTTCTGCTAGAGCTGGAATGAGAGAACAAACA 810
 QY 181 ArgilleasnaLysCysThrgLinsAlaleuMetGlnValalaspPheleuSerArg 200
 DB 811 CGGATTTATCTGAATTTGTAACAGAGCTTTGATGAGAGGTAGATTCTTCCGG 870
 QY 201 AsnLysGlnLeuTyrrGlnLysThrgLuleuSerleuGluLysProleuLeuHis 220
 DB 871 AACAAACAACCTTTATCAGAACAGCGAAATTCGTCTCAGGAGAGCCCTTCTTACAT 930
 QY 221 ThrGlyMetGlyArgleuCysThrcLinspGlnSerValSerleuAlaThrcMetlleasp 240
 DB 931 ACTGGAATGGAGCGGTTATGCACACATGATGATCTGCTCCCTGGCAACATGATTAT 990
 QY 241 ArgilleysargHisleuLysleuSerHislleArgleuAlaLeuGlyValargThr 260
 DB 991 CGAATTAACAAACACACCTAAACTATCTCATATTCCTTACGCTTGGGGGAGAAC 1050
 QY 261 LeuGlnSerGlnValLysValAlaAlaleuysAlaGlySerGlySerValleucln 280
 DB 1051 TTAGAGCTCAAGTCAAGAGTGGGCTGCTGCTGCTGGAGAGCGCTTCTTGAG 1110
 QY 281 GlyValGlnAlaAspLeuTyrrLeuThrcGlyLumetSerHisAspThrleuaspAla 300
 DB 1111 GGTGTGGAGGCTGACCTTACTCTCAGAGTGAAGTCCCATCATGATCTTGATGCT 1170
 QY 301 AlaserGlnGlylleasnaValilleuLysGlnHisSerAsnThrgluArglyPheleu 320
 DB 1171 GCTTCCCAAGAAATTAATGATCTCTGTAACACACACACAGAGAGGCTTCTT 1230
 QY 321 SeraspLeuArgaspMetleuAspSerHisleuGlnLyslleasnlleleuSer 340
 DB 1231 TCTGACCTTCGAGATATCTGATCTTCTCACTTGGAGATTAAGATTAATATCTATCA 1290
 QY 341 GluThrAspArgaspProleuGlnValal 350
 DB 1291 GAGACTGACAGGAGACCTCTCAGGTGTA 1320
 RESULT 6
 AK023378 1574 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ13316 f1s, clone OVARC1001555, weakly similar
 DEFINITION to NGG1-INTERACTING FACTOR 3.
 ACCESSION AK023378
 VERSION AK023378.1 GI:10435289
 KEYWORDS Oligo cloning; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Shiraori,A., Sudo,H.,
 Matsushima,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,

Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
 Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
 Nakamura,Y., Nagahata,K., Masuko,Y., Niomiya,K. and Iwayanagi,T.
 MEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1574)
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction; 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 source
 1. 1574
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OVARC1001555"
 /tissue_type="ovary, tumor tissue"
 /clone_lib="OVARC1"
 /note="Cloning vector: pME18SFL3"
 271. 1323
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14551.1"
 /db_xref="GI:10435290"
 /translation="MDLKALISLNDPASFASWDNVGLVPPSPHYVNTLPLIN
 DLTEVEVEVLOKADLILSYHPPIFRPMKRTIWNWKERLVIALENRGVISPDA
 YDAPOGVNMLLAGLGACTSRPIHPSKAPYPTGHNHRYEVNYYODLDKVASAVK
 GIDAVSVTSFSAFTGNEQRTIRINCOKALMOYVDFSRKOLYOTELISLKPIL
 LHTGMGRCTCTDEVSILATMTDRIKRHILKSHIILALGVGRLETSOVKYVALGSGSS
 SYLVGVENDLITGEMSHHTLDASGGINVITLDEHSNTENGFLSLDRMDLSLENK
 INILISERDRLPQV"
 BASE COUNT 420 a 361 c 372 g 421 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,24e-146 Length: 1574
 Score: 1799.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-745-506-37 (1-350) x AK023378 (1-1574)
 QY 1 MetaspLeuLysAlaLeuLeuSerSerleuAsnAspPheAlaSerleuSerPheAlaGlu 20
 DB 271 ATGATTTGAAGGCTCTCTCTTCTCTTGAATGATGATGATGATGATGATGATGATG 330
 QY 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProProHisThValAsnThreu 40
 DB 331 AGTTGGGCAATGTTGGAATCTGCTGACACCAACCCACACACATACATGTAATACACTG 390
 QY 41 PheleuThrasAspLeuThrcGlnGluValMetGlnGluValLleuGlnLysAlaAsp 60
 DB 391 TTCTGACCAATGACCTACTGAGAGATGATGAGAGGAGGCTGCAAAAGAGGAGAC 450
 QY 61 LeuileuSerTyrrHisProProlePheargProMetLysArgGileThrrTrpAsnThr 80
 DB 451 CTCATTCCTCTTACACACCGGCTATCTTCCGACCAACGAGGAGATTAACCTGGAACACA 510
 QY 81 Trrpysgluargleuvalilleargalaleuclunsnargvalglylletyrserrprohis 100
 DB 511 TGGAGGAGCGCGCTGATATCCGGGCTTCGAGAACAGAGCGGTATCTTACTCTCCAT 570
 QY 101 ThrilatrraspaalaleaProglnglyValasnsntrpleaualalyglyleuclvala 120

```

Db      571 ACAGCCTATGATGCGCCGCCAGGCGCTCAACAACATGGTGGCTAAAGGCTTGAGCT 630
Oy      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140
Db      631 TGTACCTCCAGGCCCATCATCATCTCTCCAAAGCTCCCAACCTACCCCTACACAGGAAACAC 690
Oy      141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValys 160
Db      691 CCAGTAGAATTCACAGCTTACCTACACCCCAAGCTGGACAAAGTCTGCTGCAGTGGAA 750
Oy      161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
Db      751 GGAATTGACGGTGTCTCTGCTACTCTCTTTCTGCTGACAGCTGGTAATGACAGAACAA 810
Oy      181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
Db      811 CGGATTAACTCAATATGTACTCAGAAAGCGCTTGATGACAGTGGTAAATTTCTTTCGCG 870
Oy      201 AsnLysGlnLeuTyrGlnLysThrGlnIleLeuSerLeuGluLysProLeuLeuAsnHis 220
Db      871 AACAAACAACCTTATTCAGAAAGCGAAATTCCTCTCCTGACAGAACCTTTGCTTCTACAT 930
Oy      221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
Db      931 ACTGGAAATGGACGGCTTATGACACACTGATGAAATCTGCTCCCTGGCAACCATGATTGAT 990
Oy      241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
Db      991 GCAATTAATAAGACACCTAAACATCTCATATTCCTTATGCCCCCTGGGGGGGGAACACC 1050
Oy      261 LeuGluSerGlnValLysValValAlaLeuCysAlaGlySerGlySerSerValLeuGln 280
Db      1051 TTAGAGTCTCAAGTCAAACTCGAGGCCCTGTGTGCTGGTTCGGAGCAACGCTTTCGACG 1110
Oy      281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300
Db      1111 GGTTGTGAGCGACGACCTTACCTACACAGGTGAGATCTCCCATCATCATATTGGATGCT 1170
Oy      301 AlaSerGlnGlyIleLeuValIleLeuCysGlnHisSerAsnThrLysArgGlyPheLeu 320
Db      1171 GCTTCCCAAGAAATATATGTCATCTCTGTAACAGCAACCTAAGCAGGCTTTCCT 1230
Oy      321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340
Db      1231 TCTGACCTTCGAGATATGCTGATTTCTCATCTTGCAGAAATATATATATCTATCA 1290
Oy      341 GluThrAspArgAspProLeuGlnValVal 350
Db      1291 GAGACTGACAGGAGACCTCTTCAGGTGTA 1320

RESULT 7
AB038949 1606 bp mRNA linear PRI 23-FEB-2001
LOCUS      Homo sapiens ALS2CR1 mRNA for amyotrophic lateral sclerosis 2,
DEFINITION candidate 1, complete cds.
ACCESSION  AB038949
VERSION     AB038949
KEYWORDS    AB038949.1 GI:12862477

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (sites)
Hadenno,S., Yanagisawa,Y., Skaug,J., Fichter,K., Nasir,J.,
Martindale,D., Koop,B.F., Schefer,S.W., Nicholson,D.W.,
Rouleau,G.A., Ikeda,J.-E. and Hayden,M.R.
Cloning and characterization of three novel genes, ALS2CR1,
ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis
(ALS2) critical region at chromosome 2q33-q34: candidate genes for
ALS2
JOURNAL     Genomics 71 (2), 200-213 (2001)
MEDLINE     21100893

```

```

PUBMED     11161814
REFERENCE  2 (bases 1 to 1606)
AUTHORS    Hadenno,S., Ikeda,J. and Hayden,M.R.
TITLE      Direct Submission
JOURNAL     Submitted (24-FEB-2000) Shinji Hadenno, Tokai University, The
            Institute of Medical Sciences; Bohei Imai, Isehara, Kanagawa
            259-1193, Japan (E-mail:shinji@eng.med.u-tokai.ac.jp,
            Tel:+81-463-91-5095, Fax:+81-463-91-4993)
FEATURES   Location/Qualifiers
            source
              1..1606
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /chromosome="2"
                /map="2q33"
                /note="IMAGE:32584-UniGene ID: Hs.21943"
            gene
              1..1606
                /gene="ALS2CR1"
            CDS
              286..1338
                /gene="ALS2CR1"
                /note="candidate 1"
                /codon_start=1
                /product="amyotrophic lateral sclerosis 2"
                /protein_id="BAB32499.1"
                /db_xref="GI:12862478"
                /translation="MDKALSLSLNDFASLFAESMDNGLVPEPPHTVNTLFLN
                DLEFVMEVLOKKADLLISHPPIRPMKRIWTMTKEIVTALENRGVSPHDA
                YDAAPGVNWMFLAKGLACTSRPIHRSKAPNPTBGNHREPVNTODIDKMSAVK
                GIDGVTSFSARTGNEBQTRINLCTOKLMQVDFPLSKNKOYKRTLSLEKPL
                LHTGMGRITLDESVSLATWIDIRIKHLKLSIRLALGVGRITLESQKVVYALCAGS
                SVLQGVADLYLTGEMSHDPTLDAAISQGINVILCHSNTNERGLSDLRDLJSHLENK
                INILSETRDPQLQVY"
            polyA_signal
              1566..1571
                /gene="ALS2CR1"
            BASE COUNT  438 a 365 c 375 g 428 t
            ORIGIN
            Alignment Scores:
            Pred. No.: 7,42e-146 Length: 1606
            Score: 1799.00 Matches: 350
            Percent Similarity: 100.008 Conservative: 0
            Best Local Similarity: 100.008 Mismatches: 0
            Query Match: 100.00% Indels: 0
            DB: 9 Gaps: 0
            US-09-745-506-37 (1-350) x AB038949 (1-1606)
Oy      1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
Db      286 ATGGAATTTGAAGGCTCTCTTCTTCCTTGAATGACTTTGCATCCCTCGTTGCTGAG 345
Oy      21 SerThrAspAsnValGlyLeuLeuValGluProSerProThrHisThrValAsnThrLeu 40
Db      346 AGTTGGACAAATTTGGATTATCGTGGCAACCAAGCCACACATATCTGTAATATACATC 405
Oy      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysGlyAlaAsp 60
Db      406 TTCCTGACCAATACCTGACTGAGAGATGAGAGAGAGTGGTCCAAAAGACGAC 465
Oy      61 LeuIleLeuSerTyrHisProIlePheArgProMetLysArgIleThrThrAsnThr 80
Db      466 CTGATCTCTCTCAACATCGCGCTATCTTCGACCATGAAAGGCATTAACCTGGAACACA 525
Oy      81 ThrPheGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
Db      526 TGGAAAGAGGCGCTGTGTGATCCGGCTCGGAAGAGAGTCCGTACTACTCTCCAT 565
Oy      101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnThrLeuAlaLysGlyLeuGlyAla 120
Db      586 ACAGCCTATGATGCTGCGCCCGGCGCTCAACAACATGTTGGCTAAAGGCTTGAGCT 645
Oy      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140

```


Db	646	TFAGCTCCAGGCCATACATCTCTCCAAAGCTCCCAACTACCCTACAGAGGAAACAC	705	/cell_type="hematopoietic stem/progenitor cells"
Qy	141	ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys	160	/note="isolated from a Myelodysplastic Syndrome patient"
Db	706	CGAGTGAATTCACACCTTACACACCAAGACCTGGACAAAGCATGCTGCACTGAAA	765	1..1579
Qy	161	GlyIleAspGlyAlaSerValThrSerPheSerAlaArgThrGlyAsnGluGluGlnThr	180	/gene="MDS015"
Db	766	GGAATTCAGCGTCTTCTGCACTCTTTTCTGCTAGACACTGGTAATGAGAACAAACA	825	/gene="MDS015"
Qy	181	ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSerArg	200	/codon_start=1
Db	826	CGGATTAACTGCAATGTACTCAAGAGCGTTTGATGACAGTGGATTTCTTCCCGG	885	/product="MDS015"
Qy	201	AsnLysGlnLeuLysGlnLysThrGluIleLeuSerLeuGlnLysProLeuLeuLeuHis	220	/protein_id="ANG14952.1"
Db	886	AACAAACAACCTTATATACGAAGACGGAATTTCTCTCTGAGAACACCTTTGCTTACAT	945	/db_xref="GI:10197632"
Qy	221	ThrGlyMetGlyArgLeuCysThrIleAsnAspGluSerValSerLeuAlaThrMetIleAsp	240	/translation="MDKLALSSLDNRFLSFLSFAESWDNVGLVPESPPHVYVNLFLTN
Db	946	ACTGGAATGGAGCGGTTATTCACACACTGATGATATCTGCTCCCTGGCAACCATGATGAT	1005	DLEEVMEVLQKADLILSYPIPFIRPMKRINLEHMKRELVIRALENRGILSPIT
Qy	241	ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyAlaGlyArgThr	260	AYDAAPQVGNMNLALAGLACTSRPIHPKRAPIYPEGNHREPFVNYTODLDKVMASV
Db	1006	CGAATTAAGAACACCTTAAGAACTATCTCATATTCCGTTAGCCCTTGGGGTGGGAGAAC	1065	KGIDGVSYSFSFARGINEQDRIINLNCOKALMOYVDLFSNNKOLYKRTLELSLEKL
Qy	261	LeuGluSerGlnValLysValAlaIleLeuCysAlaGlySerGlySerSerValLeuGln	280	LHTEGRLCTLDESVLSTALMTDIRKRLKISHLIALGVRTLESQVKKVVALCAGGSG
Db	1066	TTTAGAGTCTCAACTCAAAAGTCGCGGCCCTGTGTGCGGTCTTCGGAGCAGCCTTTCGAG	1125	SVYLGVEADLVLTGEMSHHDLDAASGIVNVLCEHSNTERGFLSDLRMLDSHLEN
Qy	281	GlyValGluAlaAspLeuTyrIleLeuThrGlyGluMetSerHisHisAspThrLeuAspAla	300	KIIVILSETRDRLPYV"
Db	1126	GGGTGTGAGGCTCCTACCTTACCCACAGGTGAGATGTCCTCCATCATGATCTTGGATGCT	1185	
Qy	301	AlaSerGlnGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGlyPheLeu	320	
Db	1186	GCTTCCCAAGGATTAATGATCATCTCTGTGAACACAGCACATGACGAGGCTTTCTT	1245	
Qy	321	SerAspLeuArgAspMetLeuAsnAspSerHisLeuGluAsnLysIleAsnIleLeuSer	340	
Db	1246	TCTGACCTCTGCAATATGCTGAGATCTCTCACTGGAGAAATTAATATATCTTATCA	1305	
Qy	341	GluThrAspArgAspProLeuGlnValVal	350	
Db	1306	GAGACTGACAGGAGACCTCTTCAGGTGCTA	1335	
RESULT 8				
AF182416		1579 bp	mRNA	linear
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
SOURCE				

QY 220 HistHrGlyMetGlyArgLeuGlyCysThrLeuAspGluSerValSerLeuAlaThrMetIle 239
Db 904 CATACGGAATGGAGCGGTTATGACACACTGATGATGATGCTCTCCCTGGCACCAATGATT 963
QY 240 AspArgIleLeuArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArg 259
Db 964 GATCGAATAAAGACACCTAAAGACTATCATATTCGTTAGCCCTTGCGGTGGGAGAG 1023
QY 260 ThrLeuGluSerGlnValLysValAlaLeuGlyAlaGlySerGlySerValLeu 279
Db 1024 ACCCTTAGAGTCTCAATCAAGACGCGGCCCTGTGTGCTGTGGAGCCAGCGTTCG 1083
QY 280 GlnGlyValGluAlaAspLeuArgLeuThrGlyGluMetSerHisIleAspThrLeuAsp 299
Db 1084 CAGGGGTGTGAGGCTGACCTTTACCTCACAGGAGATGTCCTCATATGATGATCTTGAT 1143
QY 300 AlaAlaSerGlnGlyIleAsnValIleLeuGlyGluHisSerAsnThrGluArgGlyPhe 319
Db 1144 GCTGCTCCCAAGGAATAATGTCATCTCTGTGACACAGCAACACTGACAGGCTTT 1203
QY 320 LeuSerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeu 339
Db 1204 CTTTCTGACCTTCGAGATATGCTGATTTCTCACTTGAGAGATAAGATAATATTATCTCA 1263
QY 340 SerGluThrAspArgAspProLeuGlnValVal 350
Db 1264 TCAGAGACTGACAGGACCCCTCTTCAGGTGGTA 1296
RESULT 9
HSM805522 1353 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens mRNA: cDNA DKFZp762L015 (from clone DKFZp762L015).
DEFINITION AL834430
ACCESSION AL834430.1 GI:21740154
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1353)
AUTHORS Blocker,H., Boecher,M., Brandt,P., Mewes,H.W., Well,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1. D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp762L015) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
FEATURES
source 1. 1353
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp762L015"
/issue_type="melanoma (Memo cell line)"
/clone_lib="762 (synonym: hmel2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
polyA_signal
polyA_site 1228..1233
BASE COUNT 457 a 283 c 280 g 333 t
ORIGIN
Alignment Scores: 1.07e-138 Length: 1353
Pred. No.: 1715.00 Matches: 332
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.33% Indels: 0
DB: 9 Gaps: 0
US-09-745-506-37 (1-350) x HSM805522 (1-1353)
QY 19 AlaGluSerThrAspAsnValGlyLeuLeuValGluProSerProHisThrValAsn 38
Db 2 GCTGAGAGTGGGACAAATGTTGATTAATCTGCTGTAACCAAGCCACCACTACTGTAAT 61
QY 39 ThrLeuPheLeuThrAsnAspLeuThrGluGluValMetGluIleValLeuGlnLys 58
Db 62 ACACCTCTTCCTGACCAATGACTGAGGAGATGATGAGAGAGGCTGCAAAAGAG 121
QY 59 AlaAspLeuIleLeuSerGlyHisProProIlePheArgProMetLysArgIleThrTrp 78
Db 122 GCAGACCTCATTTCTCCATCCATCCGCTATTTCTCCAGCCCATGGAAGCCATMACCTGG 181
QY 79 AsnThrTrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSer 98
Db 182 AACACATGGAGAGAGCGCTGGTGATCCGGCTCTGGAGAACAGATCGGTATCTACTCT 241
QY 99 ProHisThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnThrLeuAlaLysGlyLeu 118
Db 242 CCTCATACAGCCTATGATGCTGGCCCGCCAGGGGCTCAACAACAGTGGCTAAAGGGCTT 301
QY 119 GlyAlaGlyThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGly 138
Db 302 GGAGCTGTGACTCCAGGCCCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGA 361
QY 139 AsnHisArgValGluPheAsnValAsnTyrThrGluAspLeuAspLysValMetSerAla 158
Db 362 AACCCAGGATGAATTAATCAACGTTAATCAACCAAGACTGGAACAAGCAAGTCTGCA 421
QY 159 ValLysGlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlu 178
Db 422 GTGAAGAGATTAACCGGTGTTCTGTGCACCTCTTTTCTGCTAGAGCTGATAGAGAA 481
QY 179 GluThrArgIleAsnLeuAsnGlyThrGlnLysAlaLeuMetGlnValValAspPheLeu 198
Db 482 CAACACAGGATTAATGTAATTTACTACGAAAGCGTTTGATGACAGGTGGATTTTCTT 541
QY 199 SerArgAsnLysGlnLeuTyrGlnLysThrGlnIleLeuSerLeuGlnLysProLeuLeu 218
Db 542 TCCCGGAACAACAACCTTATTCAGAAAGAGGAATCTGTCACTGGAGAGCCTTGCTT 601
QY 219 LeuHisThrGlyMetGlyArgLeuGlyCysThrLeuAspGluSerValSerLeuAlaThrMet 238
Db 602 CTACATACAGGATGGAGCGGTTATGACACACTGGATGAATCTGTCTCCCGGCAACCATG 661
QY 239 IleAspArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGly 258
Db 662 ATTAGATCGAATAAAGACACCTTAACATCTCATATTTCTTACCTCCCTTGGGGGGGG 721
QY 259 ArgThrLeuGluSerGlnValLysValAlaLeuGlyAlaLeuGlySerGlySerVal 278
Db 722 AGAACCTTAGAGCTCAAGTCAAGTCGCGCCCTGTGCTGTGCTGGAGACAGGTT 781
QY 279 LeuGlnGlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisIleAspThrLeu 298
Db 782 CTCACAGGGTGTGAGCTGACCTTTACCTCACAGGAGATGTCCTCATCATGATGATCTTGG 841
QY 299 AspAlaAlaSerGlnGlyIleAsnValIleLeuGlyGluHisSerAsnThrGluArgGly 318
Db 842 GATGCTGCTCCCAAGGAATAATGTCATCTCTGTGACACAGCAACACTGAACAGGC 901
QY 319 PheLeuSerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIle 338
Db 902 TTCTTTCTGACCTTGAGATATGCTGATTTCTCACTGGAGATAAGATAATATTATC 961
QY 339 LeuSerGluThrAspArgAspProLeuGlnValVal 350

COMMENT

Cambridge, CB10 1SA, UK. E-mail enquiries: humuquery@sanger.ac.uk
 On Feb 1, 2002 this sequence version replaced q1:18135182.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/BCP/Chr1>
 RP5-925F19 is from the library RPc1-5 constructed by the group of Pierre de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCIPAC2
 IMPORTANT: This sequence is not the entire insert of clone RP5-925F19. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true left end of clone RP11-478P15 is at 53375 in this sequence. The true right end of clone RP11-393M18 is at 2000 in this sequence.

FEATURES

source
 Location/Qualifiers
 1..53374
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /clone="RP5-925F19"
 /clone_1b="RPc1-5"
 BASE COUNT 17426 a 9900 c 9927 g 18121 t
 ORIGIN

Alignment Scores:

Pred. No.: 5.58e-73 Length: 55374
 Score: 976.50 Matches: 232
 Percent Similarity: 75.21% Conservative: 38
 Best Local Similarity: 64.62% Mismatches: 71
 Query Match: 54.28% Indels: 21
 DB: 9 Gaps: 2

US-09-745-506-37 (1-350) x AL645474 (1-55374)

Oy 1 MetAplleuYsAlaleuSerLeuAAspPheAlaSerLeuSerPheAlaGlu 20
 Db 8328 ATGATTTGAAGGCTCTTCTGCTCTGAAATGACTTGCATTCCTCATTTGGTGA 8269
 Oy 21 SerTPAspAnValGlyLeuValGluProSerProHisThrValAsnThrLeu 40
 Db 8268 AGTTGTGACAAATATGATTACTGTGGGAACCAAGCATCATCTAAACACACAC 8209
 Oy 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp 60
 Db 8208 TTCACAGCAACATGACTGAGTGAAGAGATGAGAGAGTCTGTGAAGAGAGCAGAC 8149
 Oy 61 LeuLeuSerThrIleProPheProIlePheArgProMetLysArgTlleThrPAsnThr 80
 Db 8148 CTCATTTCTCTCTGCATCCACCAATTTTCAGACCTATTAAGGACATTAACCTAAGAAACC 8089
 Oy 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100

Db 8088 TGAAGAACTCCTGGTAATCCAGGCTGCGAAGCAAGAGTGTATGTATTTCTCCAC 8029
 Oy 101 ThrAlaTyrAspAlaAlaProGlnGlnGlyValAsnAspThrLeuAlaLysGlyLeuAla 120
 Db 8028 ATAGCTATGATGCTCCACCCAGGAGTGTTCACATGCTGTGCTAAAGGCTGGAGCT 7969
 Oy 121 CysThrSerArgProIleHisProSerLysAlaProAspTyrProThrGluGlyAsnHis 140
 Db 7968 TGGACCTCTAGCTGATGATGCTCTTTCGAAGCTTCCAAATGACATAGAG----- 7918
 Oy 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaVal-Ly 160
 Db 7917 -----AACCC-ACCGAAGACCTGGACAAAGTACGCTGCGAGTGA 7877
 Oy 160 scgIleAspGlyValSerValThrSerPheSerAlaValGlyLysMetGluGlnThr 180
 Db 7876 AAGAGTTGCAAGTGTCTGCTGCTTGTGCTGGGATGATGATGATGATGATGATGATG 7817
 Oy 180 ArgGluAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerAr 200
 Db 7816 ATGACACAGCTCAGTATATCTAGAGAGACTTGTGATGACAGTGTGCTTCTCCCA 7757
 Oy 200 gaAsnLysGlnLeuTyrGlnLysThrGlnLysLeuSerLeuGluLysProLeuLeuH1 220
 Db 7756 GACCATTAATATGACATCAGAGAGCTGAATATATGA-CTGGAGAACCTTGTCTTACA 7698
 Oy 220 strnGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetLeas 240
 Db 7697 TACTGCAAGGCGATGTTGATCTGATGATGATGATGATGATGATGATGATGATGATG 7638
 Oy 240 pArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgTh 260
 Db 7637 GTAAATCACAAGCCACTTAAACTATTAATATGATGCTTATGTTCTTAAGATACCGAGAC 7578
 Oy 260 rLeuGlu-SerGlnValLysValValAlaLeuCysAlaGlySerGlySerSerValLeuG 280
 Db 7577 TGAAGAGTTCTAAAGCAAGGCTGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 7518
 Oy 280 lngIValGluAlaAspLeuTyrLeu-Thr-----GlyGluMetSer 293
 Db 7517 AGAGTTAGAAAGTGAACCTTACCTTACCTGAGGTAAGGTCATCCAGGCTGGAGTCT 7458
 Oy 294 HisHisAspThrLeuAspAlaAlaSerGlnGlyIleAsnValIleLeuCysGlnHisSer 313
 Db 7457 CATCATTAATATTTGTAATGCTGCTCTTAAGAGCTAAATGCTATTCATTATGAACAC 7398
 Oy 314 AsnThrGluArgGlyPheLeuSerAspLeuArgAspMetLeuAspSerHisLeuGluAsn 333
 Db 7397 AACACGTGAATTAACACTTCTTCTGACATTTGAGATATATCTGGTCTCTTCTGAAGAGT 7338
 Oy 334 LysIleAsnThrIleLeuSerGluThrAspArgAspProLeuGlnValVal 350
 Db 7337 AAGATTAAATATATCTCT-TCAGAGACAAAGACCTCTTATATGTAGTA 7288

RESULT 12
 BDI49184 796 bp DNA linear PAT 17-JAN-2003
 LOCUS BDI49184
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BDI49184.1 GI:27854942
 VERSION JP 2002191363-A/4027.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 796)
 Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
 Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 4027 03-JUL-2002;
 JOURNAL HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)

	PN	JP 2002191363-A/4027
	PD	09-JUL-2002
	PF	08-JUL-2000 JP 2000280990
	PI	TOSHIO OTA, TAKAO ISOGAKI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SATOU,
	PI	JUNICHI NAMAI, TETSUJI OHSUKI
	PC	C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
	10,	
	PC	C12P21/02, C12J01/68//C12P21/08 G06F17/30, C12M15/00, C12N5/00 CC
		Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES		
SOURCE	FT	source
	FT	/organism='Homo sapiens'.
		location/Qualifiers
		1..796
		/mol_type='genomic DNA'
		/db_xref='taxon:9606'
BASE COUNT	192 a	204 c 202 g 195 t 3 others
ORIGIN		
Alignment Scores:		
Pred. NO.:	2.76e-62	Length: 796
Score:	826.50	Matches: 170
Percent Similarity:	96.59%	Conservative: 0
Best local Similarity:	96.59%	Mismatches: 6
Query Match:	45.94%	Indels: 3
DB:	6	Gaps: 0
US-09-745-506-37 (1-350) x BD149184 (1-796)		
OY	1	MetaspLeuysalaleuLeuSerSerleuansapPhealaSerLeuSerPhealaglu 20
Dd	271	ATGATTGTGAAGGGCTCCCTTTTCCTTGAATGACTTGACCTGCCCTCGTGGTAG 330
OY	21	SerTrpAspaanValgylLeuleuValgluproserProprohstThvalasnthieu 40
Dd	331	AGTTGGGACAAATGTTGGATTACTGCTGGAACCAAGCCACCACACTGTAATAATCACCTC 390
OY	41	PheleuthirAnaSPLeuThrGluglUvaImetGluguValLleuglnLySalaaSp 60
Dd	391	TTCTGACCAATGACCTGACCTGAGAGAAGTAGTGGAGAGAGTGTGCACAANAAGGACAGAC 450
OY	61	LeulleuSerTyRHsProProilePheargProMetlySargilEhTrpAsnThr 80
Dd	451	CTCATTTCTCTCTACCAATCCCGCTATCTTCGAGACCAATGAAGCCCATMACTGAACACCA 510
OY	81	TryplysgluArgleuVallleargAlaleuglnusnarValgylletyrsProhis 100
Dd	511	TGGAAAGAGCCCCCTGGGTGATCCGGGCTCTGGAGAACAAGTCGGGTATCTCTCTCAT 570
OY	101	ThraAtyrAsPaLaalaAProGlnGlyValInsanTrpleualalySGlyLeuGlyALA 120
Dd	571	ACAGCGTATGATGCTGGGCCCCCAAGAGGCTCAACAACATGTTGGGTAAAGGGCTTGGAGCT 630
OY	121	CysthrSerarGproIlleHisProSerLySaAProasnTyrrProthnglUGlyAsnhIs 140
Dd	631	TGTACCTCTCAGGCCCATCATCTCTCCCAAACTCCAAC-TACCCGTACAGAGGAAACAC 689
OY	141	ArgValgluPheasnaValasntTythrGlnaspLeuasplysValMetSerAlVallys 160
Dd	690	CGAAT-GAATTCAACAGTTAACTAACCAACCAAGCTGGACCAAGATCATGTCTGCANTGAAA 748
OY	161	GlytllleaspglyValSerValThrSerPheSerAlaarGthrGlyAsn 176
Dd	749	GGAAATTGACGNGTTTGTCTCTTC--TTTCTGTAGACTGGTAAT 794
RESULT 13		
LOCUS	AK114307	1328 bp mRNA linear INV 30-NOV-2002
DEFINITION	Clona intestinalis CDNA, clone:c1c1054f24, full insert sequence.	

[illegible]

Db 502 ATCGCAATTCATCAATTAATGACAAAGACTTACTGATTCATTGACAAATCTTAG 561
QY 154 LysValMetSerAlaValylsGlyLeaSerValThrSerPheSerAlaArg 173
Db 562 CAATTTGTTAAATTTGTC-----AGG 582
QY 174 ThrGlyAsnGlnThrArgIleAsnLeuAsnCysThrGlnIleValMetGln 193
Db 583 CGTACTAATGTGATGCATATAGATGATGATGATGATGATGATGATGATGATGAT 642
QY 194 ValValAspPheLeuSerArgAsnLysGlnLeuTyrGlnLysThrGlnIleLeuSerLeu 213
Db 643 ATGTGTAGAT---ATGATGCACACTTAACTGATGATGATGATGATGATGATGATGAT 699
QY 214 GlyLysProLeuLeuLeuIleThrGlyMetGlyArgLeuLysThrLeuAsnGlnSerVal 233
Db 700 GGTGAACCTCCGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
QY 234 SerLeuAlaThrMetIleAspArgIleLysArgIleLysLeuSerHisIleArgLeu 253
Db 760 ACAGTTCGACAGACTATATACAGAAATTAATAAACTGCTGCTATTCCTTACGACACTT 819
QY 254 AlaLeuGlyValGlyArgThrLeuGlnSerGlnValLysValAlaLeuLysGlnVal 273
Db 820 TCTCTGGGAGAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 879
QY 274 SerGlySerSerValLeuGlnGlyValGlnAlaAspLeuTyrLeuThrGlnGlyMetSer 293
Db 880 TCTGCTTCTAGCTTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
QY 294 HisHisAspThrLeuAlaAlaAspArgIleLysValIleLeuLysGlnHisSer 313
Db 940 CACCATGAGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999
QY 314 AsnThrGlnArgGlyLeuSerAspLeuArgAspMetLeuAspSerHisLeu---Glu 332
Db 1000 AACACAGAAAGAGGATCTTAAACAGCAAGAAATTAATAAACTGCTGATGATGATGAT 1059
QY 333 AsnLysIleAsnIleLeuSerGlnTyrAspArgAspProLeuGlnVal 349
Db 1060 TCTTTAGTTAAATCATAGTCTGATGATGATGATGATGATGATGATGATGATGAT 1110

RESULT 14
AC037455/c 170586 bp DNA linear HTG 30-AUG-2001
LOCUS Homo sapiens chromosome 02 clone RP11-422L5, WORKING DRAFT
DEFINITION
AC037455 170586 bp DNA linear HTG 30-AUG-2001
AC037455
AC037455.5 GI:9887641
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
SEQUENCE DATA
Unpublished
2 (bases 1 to 170586)
Smith,D.R.
DIRECT SUBMISSION
Submitted (09-APR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On Aug 24, 2000 this sequence version replaced gi:8569072.
COMMENT
----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: <http://www.genomecorp.com/>
Contact: gic-seqcenter@genomecorp.com
----- Project Information
Center project name: hg215
----- Summary Statistics

FEATURES
source

Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 147001 bases at least Q40
Consensus quality: 157396 bases at least Q30
Consensus quality: 159713 bases at least Q20
Insert size: 168303; sum-of-contigs
Quality coverage: 3.8x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1171: contig of 1171 bp in length
1172 1271: gap of unknown length
1272 2319: contig of 1048 bp in length
2320 2419: gap of unknown length
2420 3793: contig of 1374 bp in length
3794 3893: gap of unknown length
3894 5414: contig of 1521 bp in length
5415 5514: gap of unknown length
5515 8773: contig of 3259 bp in length
8774 8873: gap of unknown length
8874 10904: contig of 2031 bp in length
10905 11004: gap of unknown length
11005 13204: contig of 2200 bp in length
13205 13304: gap of unknown length
13305 16401: contig of 3097 bp in length
16402 18227: gap of unknown length
18228 18327: contig of 1726 bp in length
18328 22401: contig of 4074 bp in length
22402 22502: gap of unknown length
22503 25177: contig of 2676 bp in length
25178 25278: gap of unknown length
25279 29893: contig of 4615 bp in length
29894 29993: gap of unknown length
29994 33457: contig of 3465 bp in length
33458 33557: gap of unknown length
33558 37405: contig of 3848 bp in length
37406 37505: gap of unknown length
37506 41734: contig of 4229 bp in length
41735 41834: gap of unknown length
41835 48999: contig of 7165 bp in length
49000 49099: gap of unknown length
49100 56076: contig of 6977 bp in length
56077 56176: gap of unknown length
56177 63604: contig of 7428 bp in length
63605 70821: gap of unknown length
70822 70921: contig of 7117 bp in length
70922 80138: contig of 9217 bp in length
80139 80238: gap of unknown length
80239 88780: contig of 8542 bp in length
88781 88880: gap of unknown length
88881 104642: contig of 15762 bp in length
104643 104742: gap of unknown length
104743 124656: contig of 19914 bp in length
124657 143891: gap of unknown length
143892 143991: contig of 19135 bp in length
143992 170586: gap of unknown length
170587 170586: contig of 26395 bp in length.
Location/Qualifiers
1. 170586
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="02"
/clone="RP11-422L5"

```
misc_feature- /clone_11b="RPCT-11"
1. 1171
/note="assembly_name:Contig1"
misc_feature 1272. 2319
/note="assembly_name:Contig13"
misc_feature 2420. 3793
/note="assembly_name:Contig17"
misc_feature 3894. 5414
/note="assembly_name:Contig18"
misc_feature 5515. 8773
/note="assembly_name:Contig19"
misc_feature 6874. 10904
/note="assembly_name:Contig20"
/note="end:17"
misc_feature 11005. 13204
/note="assembly_name:Contig21"
misc_feature 13305. 16401
/note="assembly_name:Contig22"
misc_feature 16502. 18227
/note="assembly_name:Contig23"
misc_feature 18328. 22401
/note="assembly_name:Contig24"
misc_feature 22502. 25177
/note="assembly_name:Contig25"
misc_feature 25278. 23892
/note="assembly_name:Contig26"
misc_feature 29993. 33457
/note="assembly_name:Contig27"
misc_feature 33558. 37405
/note="assembly_name:Contig28"
misc_feature 37506. 41734
/note="assembly_name:Contig29"
misc_feature 41835. 48999
/note="assembly_name:Contig30"
misc_feature 49100. 56076
/note="assembly_name:Contig31"
/note="end:396"
misc_feature 56177. 63604
/note="assembly_name:Contig32"
misc_feature 63705. 70821
/note="assembly_name:Contig33"
misc_feature 70922. 80138
/note="assembly_name:Contig34"
misc_feature 80239. 88780
/note="assembly_name:Contig35"
misc_feature 88881. 104642
/note="assembly_name:Contig36"
misc_feature 104743. 124656
/note="assembly_name:Contig37"
misc_feature 124757. 143891
/note="assembly_name:Contig38"
misc_feature 143992. 170586
/note="assembly_name:Contig39"
BASE COUNT 50417 a 34173 c 33290 g 50259 t 2447 others
ORIGIN
Alignment Scores:
Pred. No.: 1.04e-51 Length: 170586
Score: 737.00 Matches: 172
Percent Similarity: 37.23% Conservative: 0
Best Local Similarity: 37.23% Mismatches: 1
Query Match: 40.97% Indels: 290
DB: 2 Gaps: 1
US-09-745-506-37 (1-350) x AC037455 (1-170586)
QY 1 MetaspLeuYsAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB 140069 ATGGATTGGAAGGCTCTCTCTCTCTGTAATGACTTGGCATCCCTCTGTTCTGAG 140010
QY 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProPheHisThrValAsnThrLeu 40
DB 140009 AGTTGGGCAATGTTGATTGCTGTGTGACCAAGCCACCACTACTGTAAATACACTC 139950
```

```
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetCLeuGluValLeuGlnLysLysAlaAsp 60
DB 139949 TTCTGACCAATGACCTGACCTGACGAGAGTGTGAGAGGCTGTCAGAAAAGAGCGACAG 139890
QY 61 LeuLeuLeuSerTrpHisProPheArgPheArgProMetLysArgIleThrTrpAspThr 80
DB 139889 CTCATTTCTCTCCATCCATCCGCTATCTCTCCGACCATGAGAGCGCATACCTGGAACACA 139830
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGlnAsnArgValGlyIleTrpSerProHis 100
DB 139829 TGGAGAGAGCCCTGGTGTGATCCGGGCTGTGAGAGACAGATCGGTATCTCTCCAT 139770
QY 101 ThrAlaTrpAspAlaIleArgAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeu 118
DB 139769 ACAGCCTATGATGCTGGCCGCCAGGCGCTCAACAACATGTTGCTAAAGGCT-TGGTGA 139711
QY 118 ----- 118
DB 139710 GAAAGCTCTTTCATATTTGATATTTTCCCTGCAAAATCTTGAACCTTTCAGATGCTTT 139651
QY 118 ----- 118
DB 139650 CCAGTAAAGTTTAACTGCTTTAGGTTGGGAGGCGATGATTCATTTTATATATAA 139591
QY 118 ----- 118
DB 139590 TTTTAAATATATGTTGCTGAATTTGGCAGTTAAATTCATTTGAAAAAGAACATTTT 139531
QY 118 ----- 118
DB 139530 TTATTCCTTATGAGTAAGAGAGAGAGAGAGAGAGTCTTACTGAGCTGACCTTCATGATTC 139471
QY 118 ----- 118
DB 139470 TCACCTTTGTGGCACAATGCTGTACACCGACAGCATTGAAAAATGCTTGTGAAATGAAA 139411
QY 118 ----- 118
DB 139410 GCTCAGGCTCTCCAGATATACATGATGATTAAGTCAATATAGATCAGGCTTTT 139351
QY 118 ----- 118
DB 139350 AATGATGTTTGTAGTGAATGTTTACTTATGCTTACAGCAATGAGGCAATATGATGATGGA 139291
QY 118 ----- 118
DB 139290 TTTTACCTTATTTTATGATGCTTTAGTATACATATTTGGGCTTAAAGCTTTTATGATG 139231
QY 118 ----- 118
DB 139230 ATATCTGATTAATGCTGCAAAATTAAGTATGAGCATTTGCAACCTGATGATGTTTAC 139171
QY 118 ----- 118
DB 139170 ATGCTAGATTTACCAACAGCAGTACGATCTGTCTATTTATCTGAAAAATGATGATTCAG 139111
QY 118 ----- 118
DB 139110 TCTTTCTGTCAAGTTTTCAGAGATTTTAAAGATCTTTGAAAAATATTTCTGAAATA 139051
QY 118 ----- 118
DB 139050 ATTGATTTTGTATGTCATGCAAAAGCCATGATTTATGAGGAGTGTGATGATATCTA 138991
QY 118 ----- 118
DB 138990 AGGATGTGTATGTTATGTAAGCATGAGATGATCTTTTAAAGTCAATTTCTTTCAG 138931
QY 118 ----- 118
DB 138930 AAATTTGCTATATATGATTTACTTTTAAATCTTAATCTGACCCCAAAATCTCCAT 138871
```

```

OY      119      -----GlyalacysrhrSerArgProIleHisProSerLysAla 131
Db      138870  TTCCCAACAATATTTCTCTAGAGAGCTTACCTCCAGGCCATATCTCTCCAAAGCT 138811
OY      132      ProbantYrProThgluGlyasnHisArgValGluPheasnValasnTyrThrGlnasp 151
Db      138810  CCCAACTACCCTACAGAGGAAACCCACAGTAGAATTCACGTTAACTACACCCAAAGAC 138751
OY      152      LeuAapLysValMetSerAlaValIleaspGlyValSerAlaThrSerPheSer 171
Db      138750  CTGGACAAAGTATGCTCTCATGTAAGAGAAATTGACGGTGTCTGTCTCACTCTTTTCT 138691
OY      172      AlaArg 173
Db      138690  GCTAGG 138685

RESULT 15
LOCUS      AC005037      190508 bp      DNA      linear      PRI 30-SEP-2000
DEFINITION AC005037      Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.
ACCESSION  AC005037
VERSION     AC005037.2  GI:4827310
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 190508)
AUTHORS     Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
PUBMED      9847074
REFERENCE   2 (bases 1 to 190508)
AUTHORS     Abbott,A. and Le,T.
TITLE       The sequence of Homo sapiens BAC clone RP11-469M7
JOURNAL     Unpublished
REFERENCE   3 (bases 1 to 190508)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4 (bases 1 to 190508)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (14-MAY-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5 (bases 1 to 190508)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On May 14, 1999 this sequence version replaced gi:3109089.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@wustl.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_NH0469M07

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Feensen,B., Tatenio,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletzer de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-13J8; the clone sequenced to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES

source

```

1. 190508
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-469M7"
   /clone_1lb="RP11-11"
   155..222
   /rpt_family="L2"
   982..1299
   /rpt_family="Alu"
   1502..1533
   /rpt_family="(CA)n"
   1535..1823
   /rpt_family="Alu"
   2005..2308
   /rpt_family="Alu"
   2392..2562
   /rpt_family="L1"
   3284..3315
   /rpt_family="Alu"
   3316..3338
   /rpt_family="(CAA)n"
   3339..3621
   /rpt_family="Alu"
   3650..3743
   /rpt_family="AT-rich"
   4353..4664
   /rpt_family="Alu"
   4727..5018
   /rpt_family="Alu"
   5545..5585
   /rpt_family="(CATA)n"
   5586..5873
   /rpt_family="Alu"
   5876..6007
   /rpt_family="Alu"
   6327..6713
   /rpt_family="MER21_g"
   6924..7294
   /rpt_family="MER2_type"
   8027..8049
   /rpt_family="AT-rich"
   8050..8354
   /rpt_family="Alu"
   8355..8661
   /rpt_family="Alu"

```


repeat_region	/rpt_family="Ar_rich"	8356.8684
repeat_region	/rpt_family="Alu"	9450.9752
repeat_region	/rpt_family="Alu"	10371.10680
repeat_region	/rpt_family="Alu"	10762.11038
repeat_region	/rpt_family="Alu"	11039.11109
repeat_region	/rpt_family="Alu"	11138.11410
repeat_region	/rpt_family="Alu"	11446.11663
repeat_region	/rpt_family="L1"	11664.11959
repeat_region	/rpt_family="Alu"	11960.12189
repeat_region	/rpt_family="L1"	12190.12334
repeat_region	/rpt_family="Alu"	12335.12360
repeat_region	/rpt_family="L1"	12464.12514
repeat_region	/rpt_family="L1"	12515.12822
repeat_region	/rpt_family="Alu"	12823.12862
repeat_region	/rpt_family="L1"	12857.13096
repeat_region	/rpt_family="L1"	13101.13330
repeat_region	/rpt_family="MALR"	13331.13654
repeat_region	/rpt_family="Alu"	13635.13652
repeat_region	/rpt_family="MALR"	13704.13990
repeat_region	/rpt_family="L1"	14082.14387
repeat_region	/rpt_family="Alu"	14402.14704
repeat_region	/rpt_family="Alu"	14797.15095
repeat_region	/rpt_family="Alu"	15162.15166
repeat_region	/rpt_family="Martiner"	15606.15681
repeat_region	/rpt_family="L1"	15691.15827
repeat_region	/rpt_family="MERL_type"	15661.16189
repeat_region	/rpt_family="L1"	16426.16446
repeat_region	/rpt_family="(CAT)n"	16447.16730
repeat_region	/rpt_family="Alu"	17072.17119
repeat_region	/rpt_family="L1"	17212.17504
repeat_region	/rpt_family="Alu"	17869.18139
repeat_region	/rpt_family="Alu"	18168.18464
repeat_region	/rpt_family="Alu"	18606.19155
repeat_region	/rpt_family="MALR"	19454.19767
repeat_region	/rpt_family="Alu"	19807.19949
repeat_region	/rpt_family="MERL_type"	19950.20247
repeat_region	/rpt_family="Alu"	

Alignment Scores:	1,19e-51	Length:	190508
Pred. No.:	737.00	Matches:	172
Score:	37.238	Conservative:	0
Percent Similarity:	37.238	Mismatches:	1
Best Local Similarity:	40.97%	Indels:	290
Query Match:	9	Gaps:	1
DB:			
US-09-745-506-37 (1-350) x AC005037 (1-190508)			
QY 1	MetaspleuysalaleuuserSerleusnasbphelialserleuserPhealagiu 20		
DB 122461	ATGGATTGGAAGCCTCTCTTCTCTCTGAAATGACTTGCATCCCTCTCGTTCTGAG		122402
QY 21	SerTPasAsnValGlyleuLeuValGluProSerProProHisThrValAsnThrIeu 40		
DB 122401	ACATTGGACAAATGTTGGATTACTGCTGGACCAACCCACCACTACTGTAAATACACTC		122342
QY 41	PheLeuThrAsnAspleuThrgluGluValMetGluGluValLeuGluLysLysAlaSp 60		
DB 122341	TTCTCGACCAATGACCTGACTGAGGAAGATGAGAGAGTGTCTCCAAAAGAACGACAGC		122282
QY 61	IeuIleuserSerThHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80		
DB 122281	CTCATCTCTCTTACCATCGGCTATCTTCCAGCCATGACCCATACCGCTAACCTGGAAACA		122222
QY 81	TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTrSerProHis 100		
DB 122221	TGGAAAGACCGCTGGTGTATCCGGGCTGGGAGCAACAGAGTGGTATCTACTCTCTCAT		122162
QY 101	ThrAlaTrpAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeu 118		
DB 122161	ACAGCTATGATGCTGCGCCGCCAGGGCGTCAACACTGGTGTGCTAAAGGCT-TGGTGA		122103
QY 118	-----		118
DB 122102	GAAAGCCTTTCATATTGTATATTTTCCCTGCAAACTTTGAAACTTAGATGCTTT		122043
QY 118	-----		118
DB 122042	CCAGTAAGTTTTTAACCTGTTTAGGGTTGGGAGGCGCATGATTCATTTTATATATATA		121983
QY 118	-----		118
DB 121982	TTTTAAATAATGCTGCTGAATTTGGCAGTTAAATTCGATTGGAAAAGAAACATTTT		121923
QY 118	-----		118
DB 121922	TTATCTTATGAGTAGAGAGACAGAAAGTCTTCTTACTGACGTGACCTTCATGATTC		121863
QY 118	-----		118
DB 121862	TCACCTTTGTGCGACAAATGCGCTGTCAACCCAGGCACTTGGAAAATGCTTGTGAATTGAA		121803
QY 118	-----		118
DB 121802	GCTCAGAGCCCTCCAGAAATACACTGATTGATTAAGTCAAGTAATAATAGACATCAGGCTTT		121743
QY 118	-----		118
DB 121742	AATGAATGTTTTTAGTAAGTTTACTTATGCTAGACAAATGAGGGAATGATGATGGA		121683
QY 118	-----		118
DB 121682	TTTTAGCTTATTTAGTAGAGCTTATAGTTACTCATTTTGGGCTAAAGCTTTTATGATG		121623
QY 118	-----		118
DB 121622	ATAATCGATTAAATGTGTCAATAAATAGTAGACATTGCAACCTGAATGATGATTAC		121563
QY 118	-----		118
DB 121562	ATGGCAGATTATTAACAAGCAGTAGACCAATTCTGTGATTAATCTGAAAATGATGAATCAG		121503

```
QY 118 ----- 118
Db 121502 TCTTCTGTCCAGTTTTCAGAGATTAGAAATCTTTGGAAAATAATCTGAAATA 121443
QY 118 ----- 118
Db 121442 ATTGATTTTGTATGTCCATGCCAAAGCCCTAGATTATGGGACTGGTGATGTATATCTTA 121383
QY 118 ----- 118
Db 121382 AGGATGTGTGTATGTATGTATGACATGAAATGAAATCTTTTAAAAAGTCATTCTTAG 121323
QY 118 ----- 118
Db 121322 AAATTTGCCATTAATAGTTTACTTTTAAATCTTAATCTCTCACCCCCCAAACTCCAT 121263
QY 119 -----GlyAlacysThrSerArgProIleHisPrcSerLysAla 131
Db 121262 TTTCCACAAATTTTCTCTAGAGCTTGTACTCCAGGCCCATCATCTTCCAAAGCT 121203
QY 132 ProAsnTyrProThrGluGlyAsnHisArgValGluPheAsnValAspTyrThrGlnAsp 151
Db 121202 CCCAACTACCTCTACAGAGGAAACCAACGAGTAGAATTCACGTTTACTACACCCCAAGAC 121143
QY 152 LeuAspLysValMetSerAlaValIysGlyIleAspGlyValSerValThrSerPheSer 171
Db 121142 CTGGACAAAGTCATCTCTGCACTGAAGAAGAAATGACGCTGTCTCTCCTCTTTTCT 121083
QY 172 AlaArg 173
Db 121082 GCTAGG 121077
```

Search completed: August 23, 2003, 19:24:38
Job time : 3070 secs

THIS PAGE BLANK (USPTO)